(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 October 2001 (11.10.2001)

PCT

(10) International Publication Number
WO 01/75620 A1

(51) International Patent Classification[7]: G06F 13/18, 13/40

(21) International Application Number: PCT/US00/31963

(22) International Filing Date:
21 November 2000 (21.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/541,234 3 April 2000 (03.04.2000) US

(71) Applicant: ADVANCED MICRO DEVICES, INC. [US/US]; One AMD Place, Mail Stop 68, Sunnyvale, CA 94088-3453 (US).

(72) Inventor: NOVAK, Steve; AM Haideberg 13, 01468 Boxdorf (DE).

(74) Agent: APPERLEY, Elizabeth, A.; Advanced Micro Devices, Inc., 5204 East Ben White Boulevard M/S 562, Austin, TX 78741 (US).

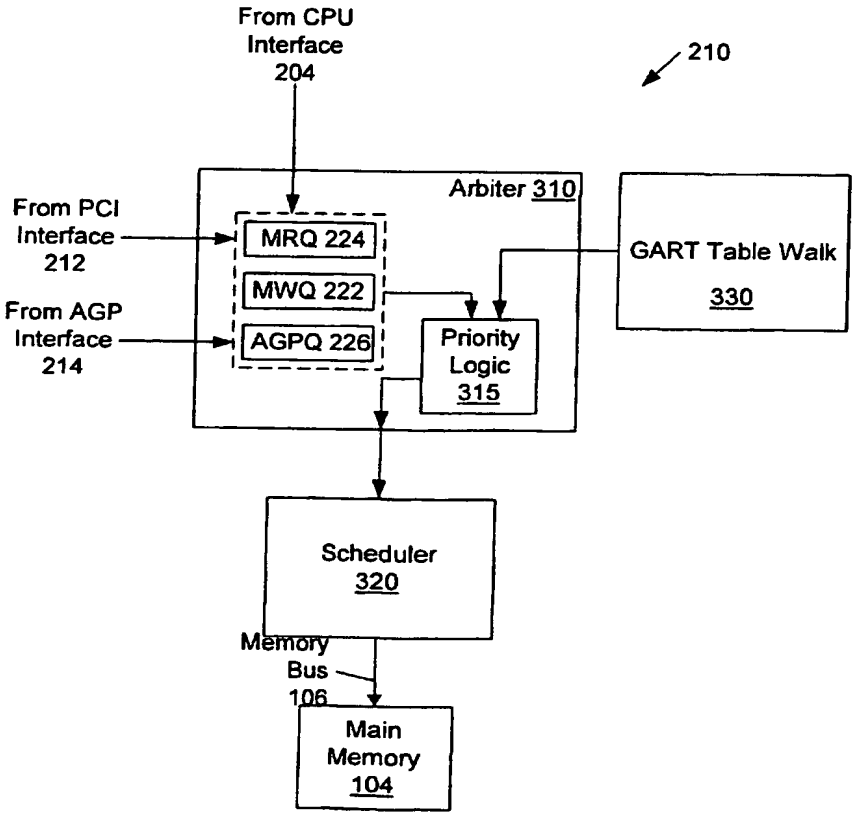
(81) Designated States *(national)*: CN, JP, KR, SG.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: BUS BRIDGE INCLUDING A MEMORY CONTROLLER HAVING AN IMPROVED MEMORY REQUEST ARBITRATION MECHANISM



(57) Abstract: A bus bridge (102) including a memory controller (210) having an improved memory request arbitration mechanism is disclosed. The memory controller (210) receives various requests to read from or write to the main memory (104). In a particular embodiment, the memory controller (210) may be configured to categorize these incoming requests into a page hit request, a page miss bank request, a page miss-different chip select request and a page conflict request. The memory controller (210) may be configured to prioritize these requests based on latency. Page hit requests have a higher arbitration priority than page miss bank requests which have a higher arbitration priority than page miss different chip-select requests which have a higher arbitration priority than page conflict requests. Since the memory controller (210) services requests based on priority, it enhances the utilization of a memory bus (106), such as an SDRAM bus.

WO 01/75620 A1

Since the page is already active, memory controller 10 initiates only a read/write cycle to read from or write to the selected page. A request to a page that is already active is commonly referred to as a page hit request. Page hit requests have the lowest latency since only one cycle is initiated by memory controller 10.

Alternatively, memory controller 10 may receive a request to open a page in a bank that has no activated page. The selected page is opened or activated by an activation cycle initiated by memory controller 10. A read/write cycle is then initiated by memory controller 10 to read from or write to the open page. A request to open a page in a bank that has no activated page is commonly referred to as a page miss bank request. Page miss bank requests have the second lowest latency because two separate cycles are initiated by memory controller 10.

Alternatively, memory controller 10 may receive a request to open a page in a different module. In what is commonly referred to as a one cycle turnaround bubble, the currently active module is first deselected and then the module that contains the page to be accessed is selected. A one cycle turnaround bubble is necessitated to ensure that the two modules, i.e. the deselected and selected modules, are not driving SDRAM bus 106 simultaneously. After the proper module is selected, the selected page is opened by an activation cycle. A read/write cycle is then initiated to read from or write to the open page. A request to open a page in a different module is commonly referred to as a page miss-different chip select request. Page miss-different chip select requests have the third lowest latency because of the one cycle turnaround bubble and the two separate cycles that are initiated by memory controller 10.

Alternatively, memory controller 10 may receive a request to open a page not active in an already active bank. The currently active page is closed or deactivated by a precharge cycle initiated by memory controller 10. The selected page is then opened or activated by an activation cycle. A separate cycle, the read/write cycle, is then initiated to read from or write to the open page. A request to open a page not active in an already active bank is commonly referred to as a page conflict request. Page conflict requests have the longest latency because three separate cycles are initiated by memory controller 10.

In some applications, memory controller 10 may be configured such that incoming requests are serviced in order of receipt. A disadvantage of memory controller 10 servicing requests in order of receipt is that a request that may have been categorized as one particular type of request may become categorized as a different type of request that has a longer latency. For example, if memory controller 10 receives a page conflict request followed by three consecutive page hit requests and services them in order of receipt, then memory controller 10 first services the page conflict request. However, after servicing the page conflict request, the page hit requests may become page conflict requests. Memory controller 10 services a page conflict request by initiating a precharge cycle to close the current page and initiating an activation cycle to activate the selected page followed by initiating a read/write cycle. However, the selected page that is opened is a different page than the page requested by the page hit requests. The page hit requests may subsequently become page conflict requests. That is, the former page hit requests may request to read from or write to a formerly active page that is no longer active. The current page may then have to be closed and a new page activated increasing the time to service the requests. A memory controller 10 servicing incoming requests in order of receipt may not optimally utilize the memory bus. Therefore it is desirable to access the SDRAM more efficiently.

Figure 1 illustrates an exemplary memory subsystem.

Figure 2 is a block diagram of one embodiment of a computer system including a bus bridge.

Figure 3 is a block diagram of one embodiment of a bus bridge.

Figure 4 is a block diagram of one embodiment of a memory controller.

Figure 5 is a block diagram of one embodiment of a priority logic block of the arbiter of Figure 4.

Figure 6 is a block diagram of one embodiment of the categorizing logic of the categorizing multiplexer of Figure 5.

Figure 7 illustrates one embodiment of the prioritizing logic of the priority select multiplexer of Figure 5.

Figure 8 illustrates one embodiment of a scheduler in a memory controller.

While the invention is susceptible to various modifications and alternative forms, specific embodiments thereof are shown by way of example in the drawings and will herein be described in detail. It should be understood, however, that the drawings and detailed description thereto are not intended to limit the invention to the particular form disclosed, but on the contrary, the invention is to cover all modifications, equivalents and alternatives falling within the spirit and scope of the present invention as defined by the appended claims.

MODE(S) FOR CARRYING OUT THE INVENTION

Figure 2 illustrates one embodiment of a computer system. Computer system 200 includes a processor 101 coupled to a variety of system components through a bus bridge 102. In computer system 200, a main memory 104 is coupled to bus bridge 102 through a memory bus 106. A graphics controller 108 is coupled to bus bridge 102 through an AGP bus 110. Furthermore, a plurality of PCI (Peripheral Component Interconnect) devices 112A-112B are coupled to bus bridge 102 through a PCI bus 114.

Processor 101 is illustrative of, for example, an Athlon™ Microprocessor. It is understood, however, that in other embodiments of computer system 200, alternative types of microprocessors may be employed. In other embodiments, an external cache unit (not shown) may further be coupled to processor 101.

Bus bridge 102 provides an interface between processor 101, main memory 104, graphics controller 108, and devices attached to PCI bus 114. When an operation is received from one of the devices connected to bus bridge 102, bus bridge 102 routes the operation to the targeted device. Bus bridge 102 generally translates an operation from the protocol used by the source device or bus to the protocol used by the target device or bus.

Main memory 104 is a memory in which application programs and data are stored. Processor 101 executes primarily out of main memory 104. In one embodiment, main memory 104 may be dynamic random access memory (DRAM) or preferably, in other embodiments, main memory 104 may be synchronous dynamic random access memory (SDRAM).

PCI devices 112A-112B are illustrative of a variety of peripheral devices such as, for example, network interface cards, video accelerators, audio cards, hard or floppy disk drives, small computer systems interface (SCSI) adapters and telephony cards. Similarly, industry standard architecture (ISA) device 118 is illustrative of one of a variety of peripheral devices such as a modem, a sound card, and a variety of data acquisition cards such as general purpose interface bus (GPIB) or field bus interface cards.

Graphics controller 108 is provided to control the rendering and images for a display device. Graphics controller 108 may embody a typical graphics accelerator generally known in the art to render three-dimensional

In this particular embodiment, main memory 104 is embodied in SDRAM. As described above, SDRAM accesses typically occur in page mode. These page mode accesses may be categorized into page hit requests (PH), page miss bank requests (PMb), page miss different chip select requests (PMc) and page conflict requests (PC). The priority given to the requests is based on the latency associated with each type of request.

Turning now to Figure 5, a block diagram of one embodiment of a priority logic block of arbiter 310 of Figure 4 is shown. Circuit components that correspond to those shown in Figure 4 are numbered identically for simplicity and clarity. Priority logic block 315 includes categorizing multiplexers 410A, 410B, 410C and 410D which use combinational logic to categorize a plurality of requests received from MRQ 224, MWQ 222 and AGPQ 226 of Figure 4 into the four categories described above: page hit (PH), page miss bank (PMB), page miss-different chip select (PMC) and page conflict (PC). For this discussion, categorizing multiplexers 410A, 410B, 410C and 410D may each be referred to as categorizing multiplexer 410. It is noted that in other embodiments, there may be additional or fewer number of categorizing multiplexers 410 depending on the number of categories. Categorizing multiplexers 410 are configured to select one request from each category, (e.g. PH, PMb, PMc and PC), based on a categorizing scheme as illustrated in Figure 6. The outputs of categorizing multiplexers 410 are input to a priority select multiplexer 420.

Priority select multiplexer 420 selects the next request to be scheduled to scheduler 320. Priority select multiplexer 420 selects from the following external/internal requests: an opportunistic precharge (OP), a high priority request from AGP interface 214 (Hi AGP), a high priority memory write request(Hi MWQ), an internal request from Gart Table Walk (GTW) 330 as well as the four categorized requests (PH, PMb, PMc, PC) from categorizing multiplexers 410, and an internal idle precharge signal. Once the requests are prioritized, priority select multiplexer 420 is configured to pass the one highest priority request to scheduler 320. The priority selection method described below in the description of Figure 7.

Priority select multiplexer 420 prioritizes the output requests from categorizing multiplexers 410 (e.g. PH's, PMb's, PMc's and PC's) based on the latency exhibited by these requests. As stated earlier, PH requests receive a higher arbitration priority over PMb, PMc and PC requests. However, to prevent starvation of service to requests with a lesser priority, priority select multiplexer 420 may be configured to stop servicing PH requests when a predetermined limit is reached. Priority select multiplexer 420 may include a programmable logic unit 421 which limits the number of page hit requests serviced consecutively from one to thirty-two. Programmable logic unit 421 includes a limit register 422 configured to store the PH request limit and a counter 423 coupled to limit register 422. Counter 423 is configured to count the number of priority hits that were serviced by scheduler 320 consecutively. Once counter 423 reaches the programmable limit, priority select multiplexer 420 may select a different request, for example, a PMb or a PMc. Counter 423 will then be reset to zero and will begin recounting the number of priority hits serviced by scheduler 320 consecutively.

Referring to Figure 6, a block diagram of one embodiment of the categorizing logic of the categorizing multiplexer of Figure 5 is shown. Circuit components that correspond to those shown in Figure 5 are numbered identically for simplicity and clarity. Categorizing multiplexer 410 uses combinational logic comprising a plurality of multiplexers to select the highest priority request. It is noted that other embodiments may use other logic blocks and configurations.

A request to initiate an OP cycle is loaded into PQ 820. A request to initiate an activation cycle is loaded into AQ 810. A request to initiate a read/write cycle is loaded into RWQ 830. It is important to note that in one particular embodiment, the size of AQ 810 and PQ 820 is one deep. The size of RWQ 830 is 3 deep. The depth of one forces scheduler to accept the next request at the latest possible time before the last request is finished processing and memory bus 106 becomes idle. Waiting until the latest possible time to accept the next request from arbiter 310 of Figure 5 gives arbiter 310 the opportunity to examine more requests, thereby increasing the probability that a PH or PMb request will be presented. It is noted however, that the size of the queues in scheduler 320 of Figure 8 may be varied in different embodiments.

In reference to the description of Figure 1, each bank of SDRAM module 16A-C includes a plurality of pages that are accessed by a particular memory access address. Depending on the particular type of request received by arbiter 310 of Figure 4, one or more cycles will be initiated by scheduler 320 of Figure 8 to access the requested page. For example, when scheduler 320 receives a PH request, a request to initiate a read/write cycle is loaded in RWQ 830 of Figure 8. Hence only a read/write cycle will be initiated by scheduler 320 to read from or write to the address location. When scheduler 320 receives a PMb request, a request to initiate both an activation cycle and a read/write cycle will be loaded in AQ 810 and RWQ 830, respectively. Since AQ 810 has a higher priority than RWQ 830, the activation cycle will be initiated before the read/write cycle. When scheduler 320 receives a PMc request, a request to initiate an activation cycle and a read/write cycle will also be loaded in activate queue 810 and read/write queue 830, respectively. After the one cycle turnaround bubble described in the description of Figure 1, scheduler 320 of Figure 8 will initiate an activation cycle followed by a read/write cycle.

If scheduler 320 receives a PC request, a request to initiate a precharge cycle, an activation cycle and a read/write cycle will be loaded in PQ 820, AQ 820 and RWQ 830 respectively. Since PQ 820 has the highest priority, the precharge cycle will execute first, then the activation cycle followed by the read/write cycle. However, since PC requests typically offer the worst latency, a PC request may be converted to an OP request followed by a PMb request. This allows more optimal requests that are received later than the PC to be interleaved between the OP and the PMb.

If scheduler 320 receives a BYP request, conceptually this would be a PMc request. However, it is treated as a PMb request since by definition there are no active chip selects. Scheduler 320 may therefore pass this request through AQ 820 to memory bus 106 in one cycle.

In addition to scheduling and optimizing events on memory bus 106, scheduler 320 may also be configured to control various other SDRAM related activities including: refresh cycles to the SDRAM, initialization and configuration of the SDRAM out of reset and power up and power down of the SDRAM.

Although the system and method of the present invention is described in connection with several embodiments, it is not intended to be limited to the specific forms set forth herein, but on the contrary, it is intended to cover such alternatives, modifications, and equivalents, as can be reasonably included within the spirit and scope of the invention as defined by the appended claims.

INDUSTRIAL APPLICABILITY

This invention is applicable to computer systems.

6. The bus bridge (102) as recited in claim 5, wherein said memory controller (210) further comprises:

a scheduler (320) coupled to receive said priority request selected by said priority select multiplexer unit (420), wherein said scheduler (320) is configured to schedule said priority request onto said synchronous dynamic random access memory bus, wherein said scheduler comprises a pre-charge queue (820), an activate queue (810), and a read/write queue (830), wherein a precharge cycle is loaded into said pre-charge queue, wherein an activation cycle is loaded into said activate queue, wherein a read/write cycle is loaded into said read/write queue.

7. The bus bridge (102) as recited in claim 1, wherein said memory controller (210) is further configured to categorize said plurality of incoming requests into a page conflict request, wherein said memory controller (210) is configured to provide said miss bank-different chip select request with a higher arbitration priority than said page conflict request.

8. The bus bridge (102) as recited in claim 7, wherein said plurality of incoming requests comprises:
at least one request from said processor bus interface (204);
at least one request from said first peripheral bus interface (212);
at least one request from said second peripheral bus interface (214); and
at least one request from a graphics address remapping table (GART) table walk (330).

9. A method of operating a memory controller (210) comprising:
receiving incoming requests from a processor bus interface (204), and a first peripheral bus interface (212);
categorizing said incoming requests into a plurality of categories, wherein said plurality of categories comprise a page hit request, a page miss bank request, and a page miss-different chip select request;
prioritizing among said incoming requests, wherein said page hit request receives a higher arbitration priority than said page miss bank request, wherein said page miss bank request receives a higher arbitration priority than said page miss-different chip select request.

10. A computer system comprising:
a microprocessor (101) coupled to a processor bus interface (204), wherein said microprocessor (101) generates one or more requests;
a main memory (104) coupled to a memory bus interface;
a bus bridge (102) coupled to provide an interface between said processor bus interface (204), said memory bus interface, a first peripheral bus interface (212), and a second peripheral bus interface (214), wherein said first peripheral bus interface (212) generates at least

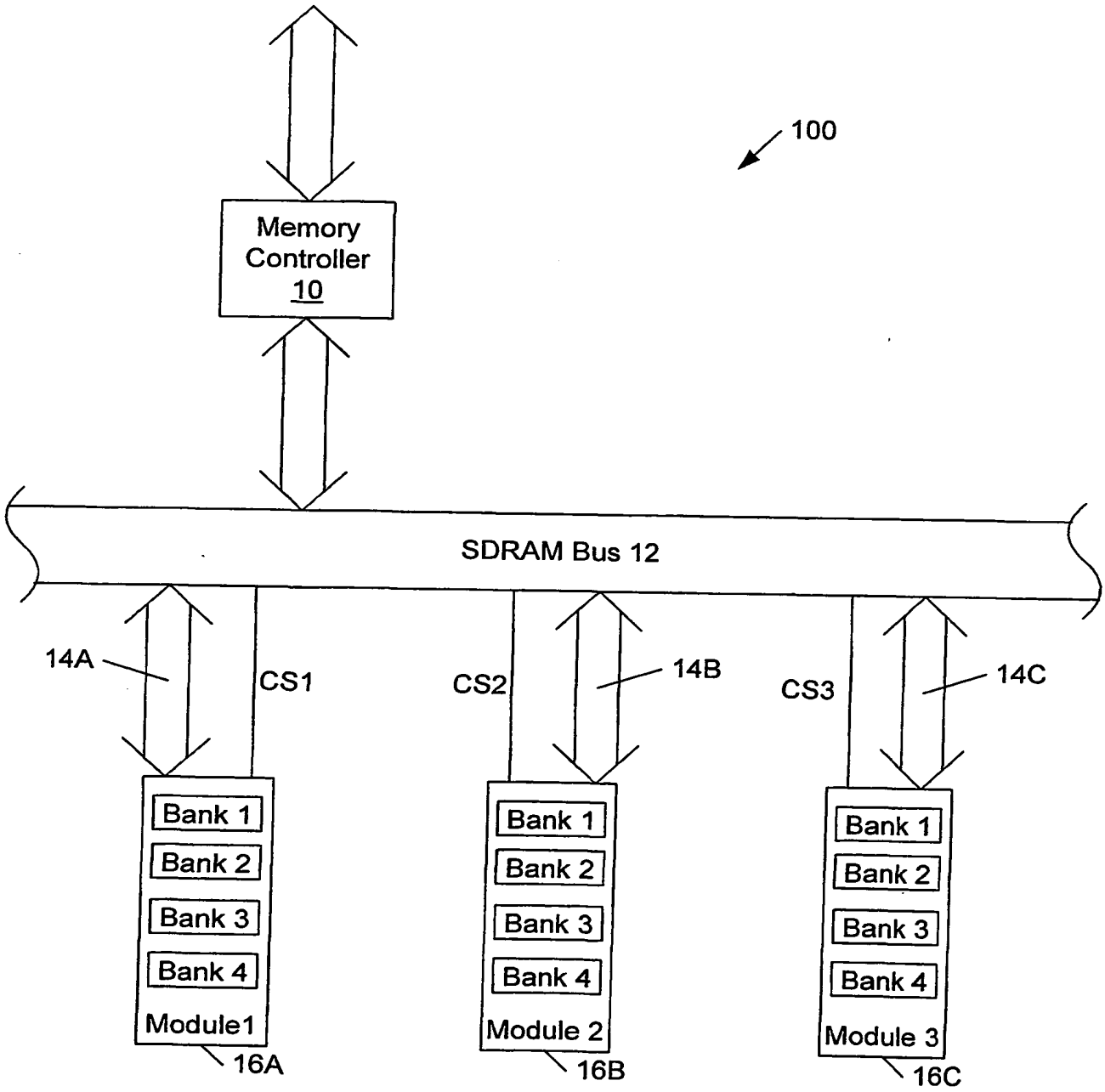


Fig. 1
(Prior Art)

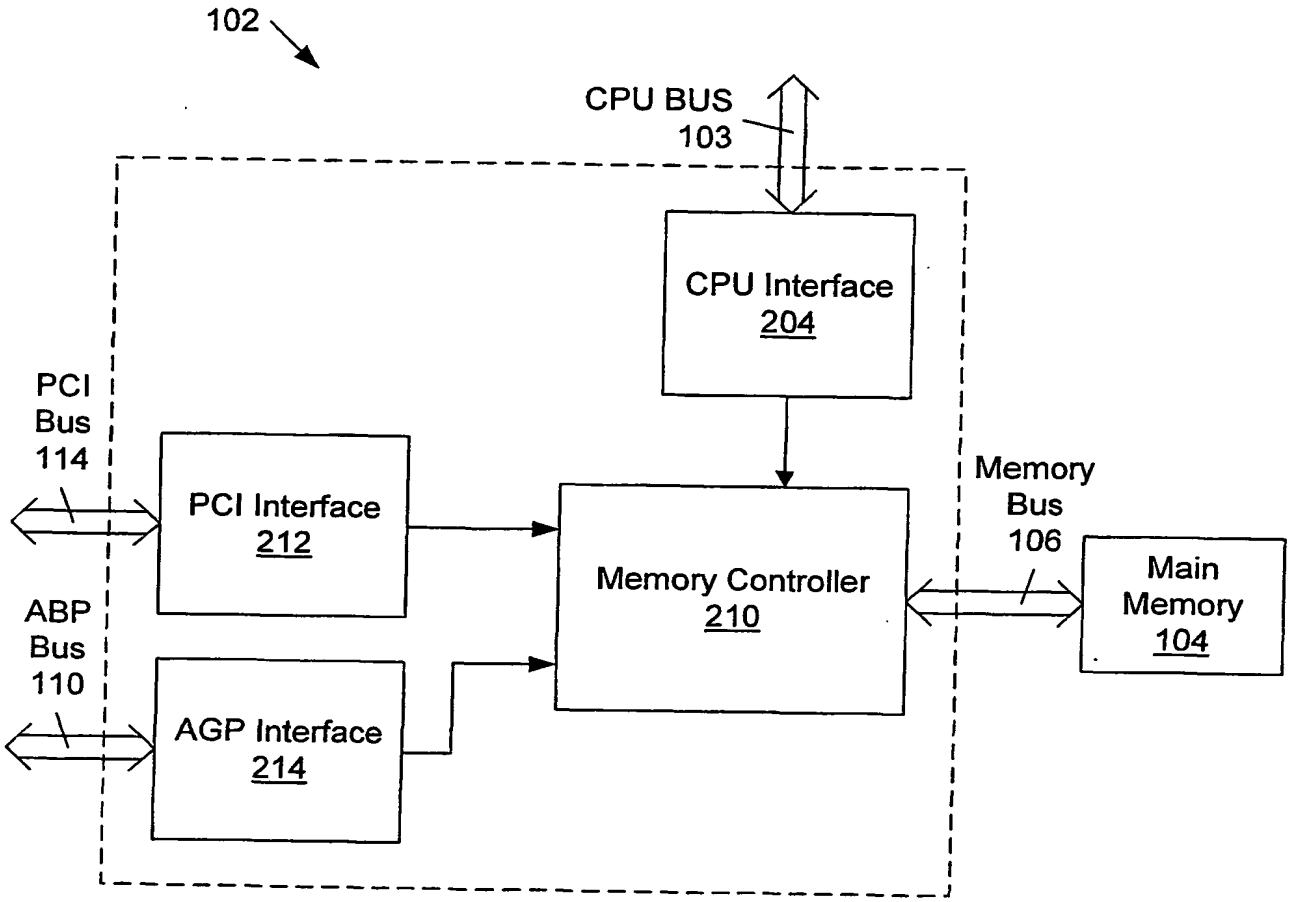

102
CPU BUS
103
CPU Interface
204
PCI
Bus
114
PCI Interface
212
Memory
Bus
106
Memory Controller
210
Main
Memory
104
ABP
Bus
110
AGP Interface
214

Fig. 3

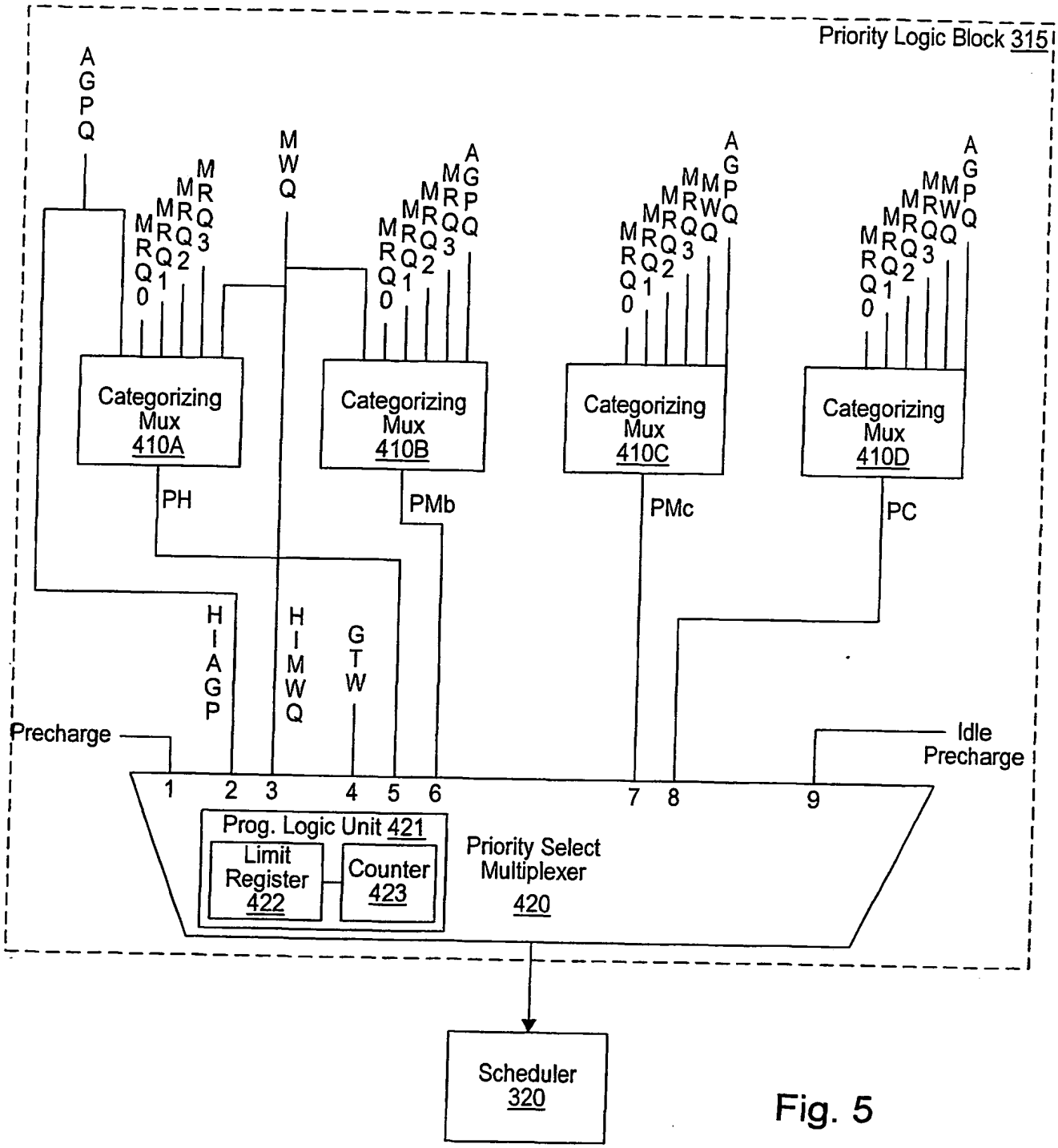


Fig. 5

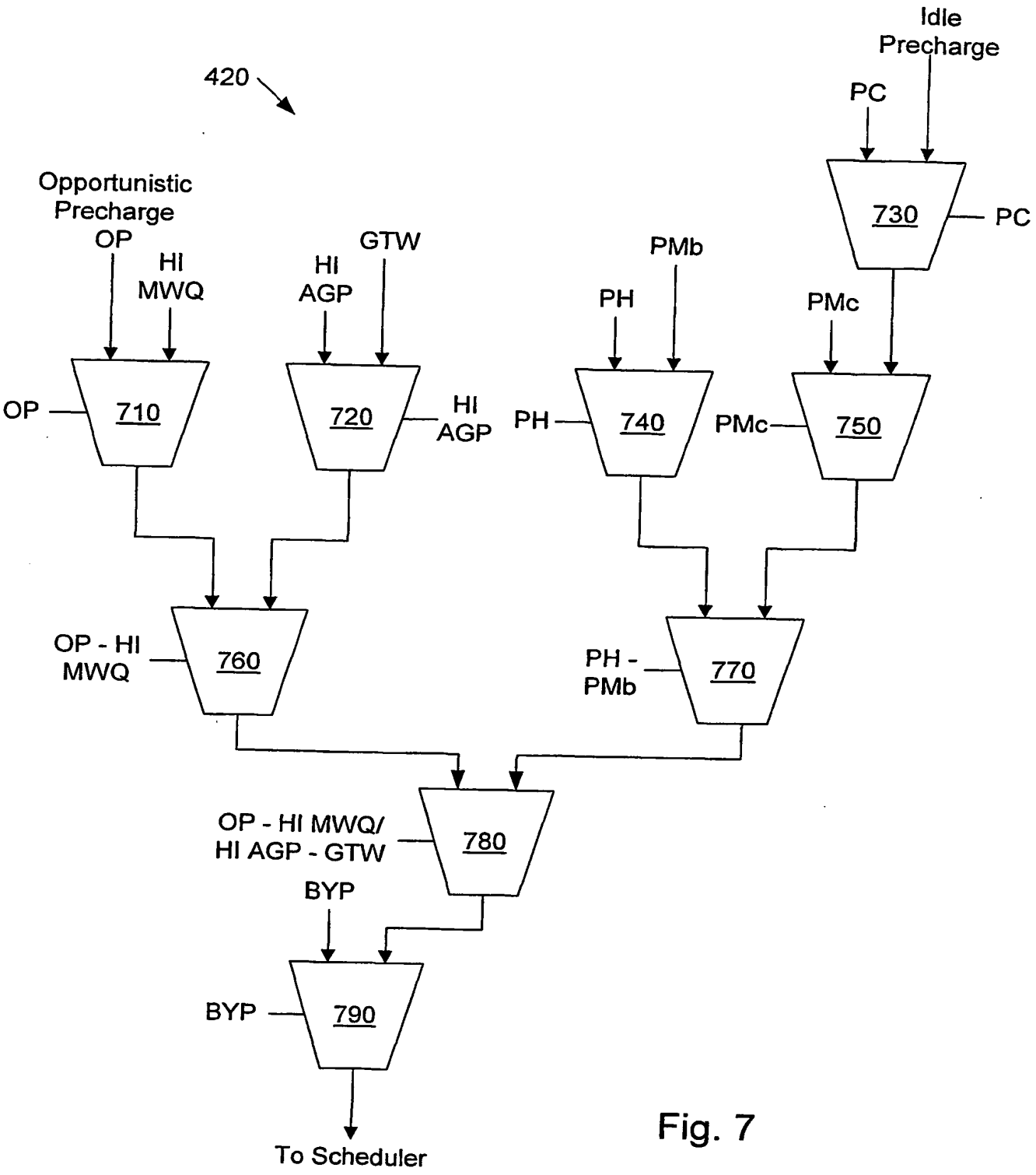

420
Idle Precharge
PC
730
PC
Opportunistic Precharge
OP
HI MWQ
HI AGP
GTW
PMb
PH
PMc
OP
710
720
HI AGP
PH
740
PMc
750
OP - HI MWQ
760
PH - PMb
770
OP - HI MWQ/ HI AGP - GTW
780
BYP
BYP
790
To Scheduler

Fig. 7

INTERNATIONAL SEARCH REPORT

Int al Application No
PCT/US 00/31963

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F13/18 G06F13/40

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, INSPEC, IBM-TDB, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 924 621 A (COMPAQ COMPUTER CORP) 23 June 1999 (1999-06-23) column 1, line 8 - line 15 column 1, line 29 - line 41 column 5, line 2 - line 8 column 2, line 50 - line 54 claim 1; figure 1	1,2,7-10
Y	US 5 745 913 A (PATTIN JAY C ET AL) 28 April 1998 (1998-04-28) column 1, line 41 -column 2, line 38 claims 1-4; figure 4	1,2,7-10
A	EP 0 391 517 A (DIGITAL EQUIPMENT CORP) 10 October 1990 (1990-10-10) column 2, line 22 - line 48 column 8, line 58 -column 9, line 35; figure 3	1-10

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 May 2001	28/05/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Henneman, P

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. Application No
PCT/US 00/31963

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0924621 A	23-06-1999	JP 11316738 A	16-11-1999
US 5745913 A	28-04-1998	NONE	
EP 0391517 A	10-10-1990	US 5222223 A	22-06-1993
		AT 156278 T	15-08-1997
		AU 628530 B	17-09-1992
		AU 5394790 A	19-12-1991
		CA 1325285 A	14-12-1993
		DE 69031139 D	04-09-1997
		DE 69031139 T	19-02-1998
		JP 2289013 A	29-11-1990
		JP 5026215 B	15-04-1993
US 5638534 A	10-06-1997	AU 693668 B	02-07-1998
		AU 5039896 A	10-10-1996
		EP 0838057 A	29-04-1998
		JP 9055081 A	25-02-1997
		SG 73403 A	20-06-2000
		WO 9630838 A	03-10-1996